Section 8.3- Error and Power

- When we reject our null hypothesis, are we definitely right? Is the null absolutely positively wrong?

NO!

- When we fail to reject our null hypothesis, is it absolutely true?

NO!

Ex: innocent vs. not guilty

Ho:

Ha:

Types of Error:

TRUTH Ho

Hyp Dec

	Ho True	Ho False
Reject Ho	Type I Error α	correct power
Fail to Reject Ho	correct	Type II Error β

Type I Error =
- reject Ho when Ho true

* more serious *

Ex: religion, business

- P(Type I Error) = α

sign. level

sex

Type II Error = (accept)
- fail to reject Ho when Ho false

less serious

- P(Type II Error) = β

POWER — good

Power = reject Ho when Ho false.

$\underbrace{\text{when Ho false}}_{\text{another prop is true}}$

- Formula:
  - prob.
  - Power = $1 - \beta$

$\beta = 1 - \text{power}$

$H_0: p = 0.30$

$H_a: p > 0.30$

$p = 0.70$

$p = 0.05$

Calculating Power

prob. of rejecting Ho when p_A is true

1. STATE: important info
   - Ho, Ha
   - n
   - α
   - p_A = alternative prop. that's true

$H_0: p = 0.30$

$H_a: p > 0.30$

$\hat{p} = 0.70$

$\hat{p} = 0.60$

$\hat{p} = 0.35$

0.32

$\hat{p} = 0.05$

2. FIND:
   - rejection region = values of $\hat{p}$ that will reject Ho
     * bypass z & p-val.
   - need: α & type of test (1 sided or 2-sided)

3. CALCULATE:

- prob. $\underbrace{\text{of getting rejection region}}_{\text{rejecting } H_o}$ $\underbrace{\text{if } p_A \text{ is true}}_{\text{when } H_o \text{ is false}}$

4. Power is considered adequate when:

   $> 80\%$

5. This is only done on...

   1 prop. tests

   $H_o: p =$

   $H_a: p \substack{> \\ < \\ \neq}$

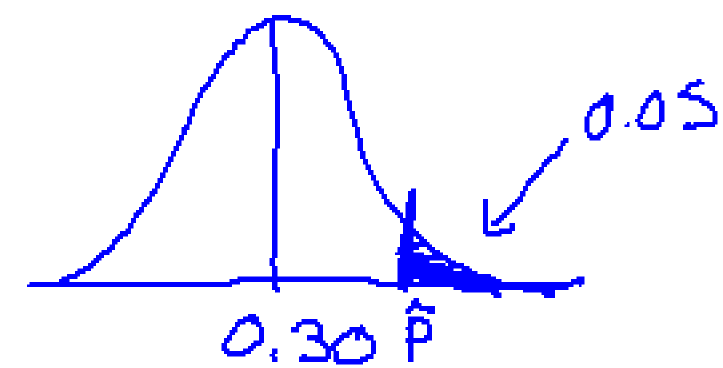
__Example:__ We want to examine the color distribution in a bag of Skittles. We believe that the machine that fills these bags is not putting the right amount of colors in each because we have gotten 3 bags recently that have a lot of red in them. There are claimed to be 30% reds in each bag. To check whether the colors are correct, employees select a random sample of 35 packages. Is this test sufficiently sensitive to detect an increase of 0.05 in the proportion of red Skittles? Calculate the power of this test against the alternative.

State the hypotheses, the alternative that we want to detect and the significance level.

$H_0: p = 0.30$

$H_a: p > 0.30$ $\quad n = 35 \quad \alpha = 0.05 \quad p_A = 0.35$

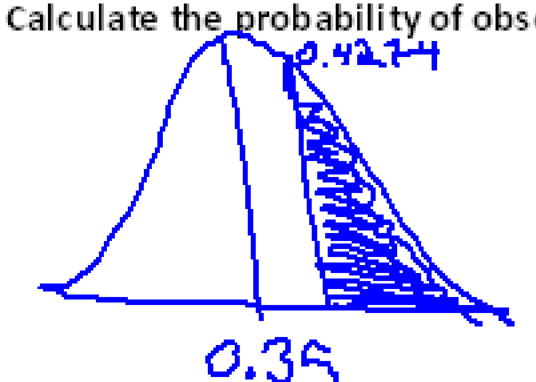
Find the values of p that will lead to the rejection of Ho. (rejection region)



$$\hat{p} = \text{invnorm}\left(0.95, 0.30, \sqrt{\frac{(0.3)(0.7)}{35}}\right)$$

$$\hat{p} = 0.4274$$

Calculate the probability of observing these values of **p** when the alternative is true.



$$P(\hat{p} > 0.4274 \mid p = 0.35)$$

$$= \text{normalcdf}(0.4274, e99, 0.35, \sqrt{\frac{(0.35)(0.65)}{35}})$$

$$= 0.1588$$

Find the probability of Type I and Type II errors.

Type I error $= \alpha = 0.05$

Type II error $= \beta = 1 - 0.1588 = 0.8412$

Example: (2 sided test with power)

Ho: p = 0.75

Ha: p $\neq$ 0.75

alpha = ~~0.01~~ $\rightarrow$ 0.05

~~n = 60~~

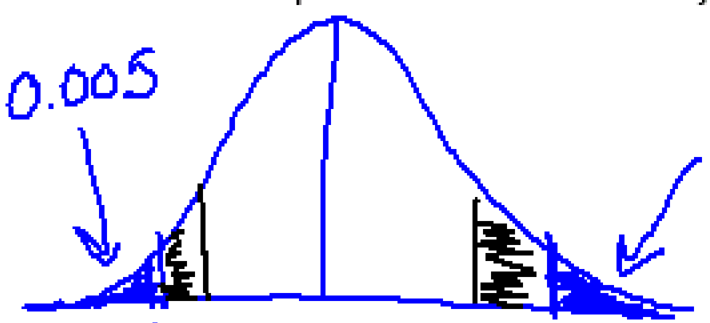
What is the power of the test against the alternative of p = 0.80? $p_A = 0.80$

$\hat{p} = 0.95$ $\hat{p} = 0.35$

State the hypotheses, the alternative that we want to detect and the significance level.

Stated above

Find the values of p that will lead to the rejection of Ho. (rejection region)



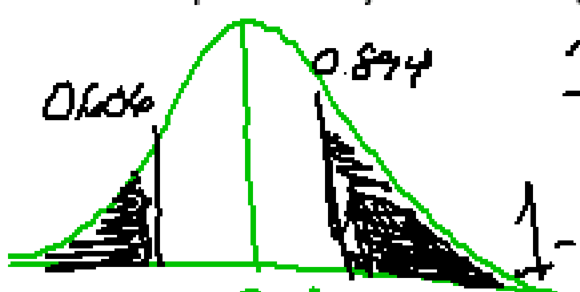
$$\hat{p}_1 = \text{invnorm}\left(0.005, 0.75, \sqrt{\frac{(0.75)(0.25)}{60}}\right)$$

$$\hat{p}_1 = 0.606$$

$$\hat{p}_2 = \text{invnorm}\left(0.995, 0.75, \sqrt{\quad}\right)$$

$$\hat{p}_2 = 0.894$$

Calculate the probability of observing these values of $\hat{p}$ when the alternative is true.



$$1 - \text{normalcdf}\left(0.606, 0.894, 0.80, \sqrt{\quad}\right)$$

$$1 - P(0.606 \leq \hat{p} < 0.894 \mid p = 0.80) = 0.0466$$

No!

Find the probability of Type I and Type II errors.

Type I = 0.01

Type II = 1 − 0.0466 ≈ 95%

Increasing Power

3 things we can do to increase power are….

1. Increase n

- more data $\Rightarrow$ more info. about pop.

2. Increase α

- bigger rejection region $\Rightarrow$ bigger power
- easier to reject @ $\alpha = 0.10$ than 0.01

3. Consider

p_a further from p

$H_0: p = 0.30$

Ex: wt. gain

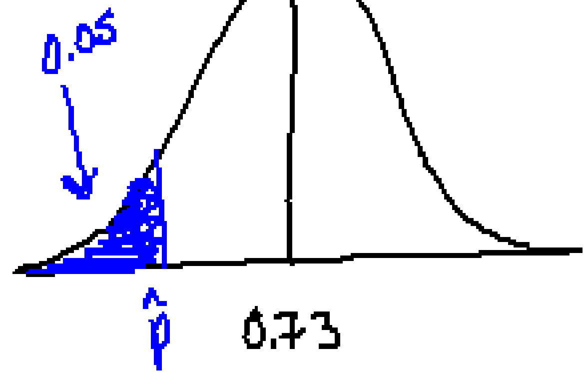
① $n = 100$

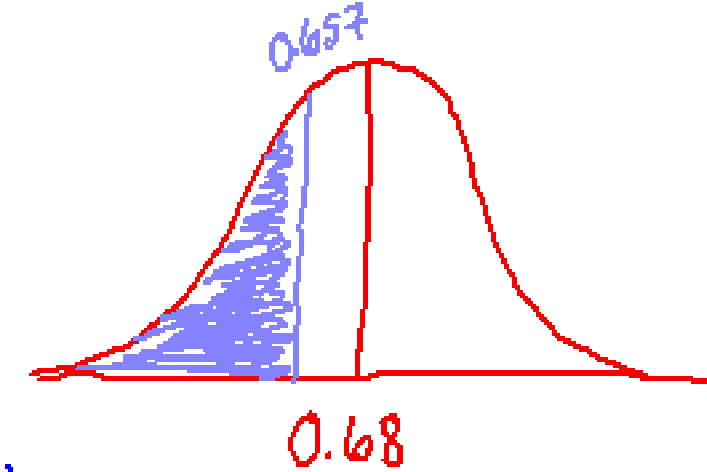
$H_0: p = 0.73$

$H_a: p < 0.73$

$p_A = 0.68$

$\alpha = 0.05$





$\hat{p} = \text{invnorm}\left(0.05, 0.73, \sqrt{\frac{(0.73)(0.27)}{100}}\right)$

$= 0.657$

$P(\hat{p} < 0.657 \mid p = 0.68)$

$\boxed{= 0.3022}$ = power

Type I = 0.05

Type II = 0.6978

② $n = 125$

$H_0: p = 0.10$

$H_a: p > 0.10$

$\hat{p} = 0.13$

$\alpha = 0.05$

$\hat{p} = \text{invnorm}(0.95, 0.10, \sqrt{\quad}) = 0.1441$

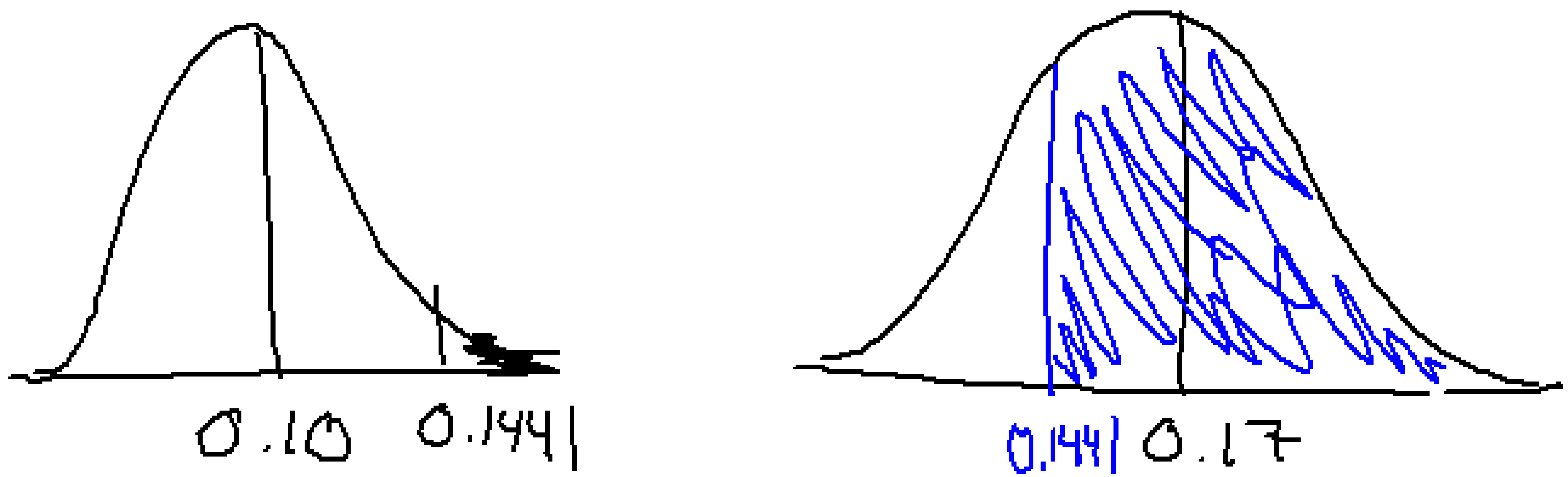
$P(\hat{p} > 0.1441 \mid p = 0.13)$

$= 0.2996$

Type I = α = 0.05

Type II = β = 0.7004

③ in prob. #2, change:
Will the test sufficiently detect a change in 7% ?



power = 0.8328

yes power > 80%